



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ : C12N 15/12, C07K 14/47, C12Q 1/68, A61K 38/17, C07K 16/18	A1	(11) International Publication Number: WO 00/29580 (43) International Publication Date: 25 May 2000 (25.05.00)
---	----	---

(21) International Application Number: PCT/US99/27027

(22) International Filing Date: 12 November 1999 (12.11.99)

(30) Priority Data:
09/190,965 13 November 1998 (13.11.98) US(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Application
US 09/190,965 (CIP)
Filed on 13 November 1998 (13.11.98)

(71) Applicant (for all designated States except US): INCYTE PHARMACEUTICALS, INC. [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): TANG, Y., Tom [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). GUEGLER, Karl, J. [CH/US]; 1048 Oakland Avenue, Menlo Park, CA 94025 (US). CORLEY, Neil, C. [US/US]; 1240 Dale Avenue, #30, Mountain View, CA 94040 (US). GORGONE, Gina, A. [US/US]; 1253 Pinecrest Drive, Boulder Creek, CA 95006 (US).

(74) Agents: BILLINGS, Lucy, J. et al.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: CALCIUM BINDING PROTEIN

```

          9          18          27          36          45          54
5' CAG AAA CAG AAC TGC CTG TGA CAG ATT AAG AGA CAA GCA AGG CTT GGA ATC TGA

          63          72          81          90          99          108
GAG CAA GCA AAG AGA GTG GAA ATT TAC AGC TGC CTT ATC ATT CCA TAT TGG AAG

          117          126          135          144          153          162
AAG AGA TTT CTA CAC ATG AAA AAA ATG CCT TTG TTT AGT AAA TCA CAC AAA AAT
              M   K   K   M   P   L   F   S   K   S   H   K   N

          171          180          189          198          207          216
CCA GCA GAA ATT GTG AAA ATC CTG AAA GAC AAT TTG GCC ATT TTG GAA AAG CAA
P   A   E   I   V   K   I   L   K   D   N   L   A   I   L   E   K   Q

          225          234          243          252          261          270
GAC AAA AAG ACA GAC AAG GCT TCA GAA GAA GTG TCT AAA TCA CTG CAA GCA ATG
D   K   K   T   D   K   A   S   E   E   V   S   K   S   L   Q   A   M

          279          288          297          306          315          324
AAA GAA ATT CTG TGT GGT ACA AAC GAG AAA GAA CCC CCG ACA GAA GCA GTG GCT
K   E   I   L   C   G   T   N   E   K   E   P   P   T   E   A   V   A

          333          342          351          360          369          378
CAG CTA GCA CAA GAA CTC TAC AGC AGT GGC CTG CTG GTG ACA CTG ATA GCT GAC
Q   L   A   Q   E   L   Y   S   S   G   L   L   V   T   L   I   A   D

```

(57) Abstract

The invention provides a human calcium binding protein (hCBP) and polynucleotides which identify and encode hCBP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with expression of hCBP.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakistan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

CALCIUM BINDING PROTEIN

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of a calcium binding protein
5 and to the use of these sequences in the diagnosis, treatment, and prevention of cancer, reproductive disorders, immune disorders, and developmental disorders.

BACKGROUND OF THE INVENTION

In nearly all eukaryotic cells, calcium (Ca^{2+}) functions as an intracellular signaling molecule
10 in diverse cellular processes including cell proliferation, neurotransmitter secretion, glycogen metabolism, and skeletal muscle contraction. Within a resting cell, the concentration of Ca^{2+} in the cytosol is extremely low, $<10^{-7}$ M. However, when the cell is stimulated by an external signal, such as a neural impulse or a growth factor, the cytosolic concentration of Ca^{2+} increases by about 50-fold. This influx of Ca^{2+} is caused by the opening of plasma membrane Ca^{2+} channels and the release of
15 Ca^{2+} from intracellular stores such as the endoplasmic reticulum. Ca^{2+} directly activates regulatory enzymes, such as protein kinase C, which trigger signal transduction pathways. Ca^{2+} also binds to specific Ca^{2+} -binding proteins (CBPs) such as calmodulin (CaM) which then activate multiple target proteins including enzymes, membrane transport pumps, and ion channels. CaM is the most widely distributed and the most common mediator of calcium effects and appears to be the primary sensor of
20 Ca^{2+} changes in eukaryotic cells. The binding of Ca^{2+} to CaM induces marked conformational changes in the protein permitting interaction with, and regulation of, over 100 different proteins. CaM interactions are involved in a multitude of cellular processes including, but not limited to, gene regulation, DNA synthesis, cell cycle progression, mitosis, cytokinesis, cytoskeletal organization, muscle contraction, signal transduction, ion homeostasis, exocytosis, and metabolic regulation (Celio,
25 M.R. et al. (1996) Guidebook to Calcium-binding Proteins, Oxford University Press, Oxford, UK, pp. 15-20).

A novel mouse gene named MO25, expressed during early stages of development, has recently been identified and is believed to encode a CBP. The Drosophila equivalent of MO25, DMO25, encodes a polypeptide of 339 amino acid residues with a calculated molecular mass of 39.3
30 kDa. This novel CBP was found to be conserved among Drosophila, mouse, and yeast. In particular, the carboxy-terminal region of the protein is highly conserved among these species. A homology search revealed that the amino acid sequence of MO25 and DMO25 is similar to a protein encoded in an open reading frame near the calcineurin B subunit gene on chromosome XI in Saccharomyces cerevisiae. Calcineurin B is the small Ca^{2+} -binding regulatory subunit of calcineurin, a CaM-

regulated protein phosphatase. The conservation of the MO25 and DMO25 gene structure among species and the wide tissue expression profile indicates that the function of the gene is likely to be fundamental in many cell types as well as during development (Nozaki, M. et al. (1996) DNA Cell Biol. 15:505-509; and Miyamoto, H. et al. (1993) Mol. Reprod. Dev. 34:1-7).

5 CBPs are implicated in a variety of disorders. For example, calcineurin is found in the cells of all eukaryotes ranging from yeast to mammals. Calcineurin is a target for inhibition by the immunosuppressive agents cyclosporin A and FK506 emphasizing its importance in immune disorders (Kissinger, C.R. et al. (1995) Nature 378:641-644). Calcineurin also plays a critical role in transcriptional regulation and growth control in T-lymphocytes (Wang, M.G. et al. (1996) Cytogenet. 10 Cell Genet. 72:236-241). Additionally, levels of CaM are increased several-fold in tumors and tumor-derived cell lines for various types of cancer (Rasmussen, C.D. and A.R. Means (1989) Trends Neurosci. 12:433-438). Calcium binding S100 β is another example of a CBP involved in a variety of disorders. S100 β contains an EF-hand motif and is abundantly expressed in the nervous system. S100 β levels are increased in the blood and cerebrospinal fluid of patients with neurological injury 15 resulting from cerebral infarction, transient ischemic attacks, hemorrhagia, head trauma, and Down's syndrome. Furthermore, S100 β and other neural-specific CBPs may also protect against neurodegenerative disorders, such as Alzheimer's, Parkinson's, and Huntington's diseases.

The discovery of a new calcium binding protein and the polynucleotides encoding it satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and 20 treatment of cancer, reproductive disorders, immune disorders, and developmental disorders.

SUMMARY OF THE INVENTION

The invention features substantially purified polypeptides, calcium binding protein, referred to as "hCBP." In one aspect, the invention provides a substantially purified polypeptide comprising 25 an amino acid sequence selected from the group consisting of hCBP and fragments thereof. The invention also includes a polypeptide comprising an amino acid sequence that differs by one or more conservative amino acid substitutions from the amino acid sequence of SEQ ID NO:1.

The invention further provides a substantially purified variant having at least 90% amino acid identity to at least one of the amino acid sequences selected from the group consisting of SEQ ID 30 NO:1 and fragments thereof. The invention also provides an isolated and purified polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1 and fragments thereof. The invention also includes an isolated and purified polynucleotide variant having at least 70% polynucleotide sequence identity to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of

SEQ ID NO:1 and fragments thereof.

Additionally, the invention provides an isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1 and fragments thereof. The invention
5 also provides an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide encoding the polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NO:1 and fragments thereof.

The invention also provides a method for detecting a polynucleotide in a sample containing nucleic acids, the method comprising the steps of: (a) hybridizing the complement of the
10 polynucleotide sequence to at least one of the polynucleotides of the sample, thereby forming a hybridization complex; and (b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of a polynucleotide in the sample. In one aspect, the method further comprises amplifying the polynucleotide prior to hybridization.

The invention also provides an isolated and purified polynucleotide comprising a
15 polynucleotide sequence selected from the group consisting of SEQ ID NO:2 and fragments thereof. The invention further provides an isolated and purified polynucleotide variant having at least 70% polynucleotide sequence identity to the polynucleotide sequence selected from the group consisting of SEQ ID NO:2 and fragments thereof. The invention also provides an isolated and purified
20 polynucleotide having a sequence which is complementary to the polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:2 and fragments thereof.

The invention further provides an expression vector containing at least a fragment of the polynucleotide encoding the polypeptide comprising the amino acid sequence of SEQ ID NO:1. In another aspect, the expression vector is contained within a host cell.

The invention also provides a method for producing a polypeptide, the method comprising the
25 steps of: (a) culturing the host cell containing an expression vector containing a polynucleotide of the invention under conditions suitable for the expression of the polypeptide; and (b) recovering the polypeptide from the host cell culture.

The invention also provides a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1 and
30 fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention further includes a purified antibody which binds to a polypeptide selected from the group consisting of SEQ ID NO:1 and fragments thereof. The invention also provides a purified agonist and a purified antagonist to the polypeptide.

The invention also provides a method for treating or preventing a disorder associated with

decreased expression or activity of hCBP, the method comprising administering to a subject in need of such treatment an effective amount of a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1 and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

5 The invention also provides a method for treating or preventing a disorder associated with increased expression or activity of hCBP, the method comprising administering to a subject in need of such treatment an effective amount of an antagonist of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1 and fragments thereof.

10 BRIEF DESCRIPTION OF THE FIGURES AND TABLE

Figures 1A, 1B, 1C, and 1D show the amino acid sequence (SEQ ID NO:1) and nucleic acid sequence (SEQ ID NO:2) of hCBP. The alignment was produced using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA).

15 Figures 2A and 2B show the amino acid sequence alignment between hCBP (Incyte Clone ID 3734805; SEQ ID NO:1), MO25 (GI 262934; SEQ ID NO:3), DMO25 (GI 1794137; SEQ ID NO:4), and yeast-like CBP (GI 1255838; SEQ ID NO:5) produced using the multisequence alignment program of LASERGENE software (DNASTAR, Madison WI).

Table 1 shows the tools, programs, and algorithms used to analyze hCBP, along with applicable descriptions, references, and threshold parameters.

20 DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing 25 particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a 30 reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be

used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

"hCBP" refers to the amino acid sequences of substantially purified hCBP obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which intensifies or mimics the biological activity of hCBP. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of hCBP either by directly interacting with hCBP or by acting on components of the biological pathway in which hCBP participates.

An "allelic variant" is an alternative form of the gene encoding hCBP. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding hCBP include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as hCBP or a polypeptide with at least one functional characteristic of hCBP. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding hCBP, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding hCBP. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent hCBP. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of hCBP is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine.

Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

The terms "amino acid" and "amino acid sequence" refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic
5 molecules. Where "amino acid sequence" is recited to refer to an amino acid sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid sequence.

10 Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term "antagonist" refers to a molecule which inhibits or attenuates the biological activity of hCBP. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of hCBP either by
15 directly interacting with hCBP or by acting on components of the biological pathway in which hCBP participates.

The term "antibody" refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, F(ab')₂, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind hCBP polypeptides can be prepared using intact polypeptides or using fragments
20 containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

25 The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen
30 used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition containing a nucleic acid sequence which is complementary to the "sense" strand of a specific nucleic acid sequence. Antisense molecules may be produced by any method including synthesis or transcription. Once introduced into a cell, the complementary nucleotides combine with natural sequences produced by the cell to form duplexes

and to block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" refers to the capability of the natural, recombinant, or synthetic hCBP, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The terms "complementary" and "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence "5' A-G-T 3'" bonds to the complementary sequence "3' T-C-A 5'." Complementarity between two single-stranded molecules may be "partial," such that only some of the nucleic acids bind, or it may be "complete," such that total complementarity exists between the single stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands. This is of particular importance in amplification reactions, which depend upon binding between nucleic acid strands, and in the design and use of peptide nucleic acid (PNA) molecules.

A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding hCBP or fragments of hCBP may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been resequenced to resolve uncalled bases, extended using the XL-PCR kit (Perkin-Elmer, Norwalk CT) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from the overlapping sequences of one or more Incyte Clones and, in some cases, one or more public domain ESTs, using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison WI). Some sequences have been both extended and assembled to produce the consensus sequence.

"Conservative amino acid substitutions" are those substitutions that, when made, least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

	Original Residue	Conservative Substitution
	Ala	Gly, Ser
	Arg	His, Lys
	Asn	Asp, Gln, His
5	Asp	Asn, Glu
	Cys	Ala, Ser
	Gln	Asn, Glu, His
	Glu	Asp, Gln, His
	Gly	Ala
10	His	Asn, Arg, Gln, Glu
	Ile	Leu, Val
	Leu	Ile, Val
	Lys	Arg, Gln, Glu
	Met	Leu, Ile
15	Phe	His, Met, Leu, Trp, Tyr
	Ser	Cys, Thr
	Thr	Ser, Val
	Trp	Phe, Tyr
	Tyr	His, Phe, Trp
20	Val	Ile, Leu, Thr

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to the chemical modification of a polypeptide sequence, or a polynucleotide sequence. Chemical modifications of a polynucleotide sequence can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

A "fragment" is a unique portion of hCBP or the polynucleotide encoding hCBP which is identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues

in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50% of a polypeptide) as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

A fragment of SEQ ID NO:2 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:2, for example, as distinct from any other sequence in the same genome. A fragment of SEQ ID NO:2 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:2 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:2 and the region of SEQ ID NO:2 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1 is encoded by a fragment of SEQ ID NO:2. A fragment of SEQ ID NO:1 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1. For example, a fragment of SEQ ID NO:1 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1. The precise length of a fragment of SEQ ID NO:1 and the region of SEQ ID NO:1 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

The term "similarity" refers to a degree of complementarity. There may be partial similarity or complete similarity. The word "identity" may substitute for the word "similarity." A partially complementary sequence that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid is referred to as "substantially similar." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or northern blot, solution hybridization, and the like) under conditions of reduced stringency. A substantially similar sequence or hybridization probe will compete for and inhibit the binding of a completely similar (identical) sequence to the target sequence under conditions of reduced stringency. This is not to say that conditions of reduced stringency are such that non-specific binding is permitted, as reduced stringency conditions require that the binding of two sequences to one another be a specific (i.e., a selective) interaction. The absence of non-specific binding may be tested by the use of a second target sequence which lacks even a partial degree of complementarity (e.g., less than about 30% similarity or identity). In the absence of non-specific binding, the substantially similar sequence or probe will not hybridize to the second non-complementary target sequence.

The phrases "percent identity" and "% identity," as applied to polynucleotide sequences, refer

to the percentage of residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

5 Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS
10 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polynucleotide sequence pairs.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms
15 is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at <http://www.ncbi.nlm.nih.gov/BLAST/>. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other
20 polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at <http://www.ncbi.nlm.nih.gov/gorf/bl2.html>. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to
25 compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Reward for match: 1

Penalty for mismatch: -2

30 *Open Gap: 5 and Extension Gap: 2 penalties*

Gap x drop-off: 50

Expect: 10

Word Size: 11

Filter: on

Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous
5 nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes
10 in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to the percentage of residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some
15 alignment methods take into account conservative amino acid substitutions. Such conservative substitutions, explained in more detail above, generally preserve the hydrophobicity and acidity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide.

Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e
20 sequence alignment program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default residue weight table. As with polynucleotide alignments, the percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polypeptide sequence pairs.

25 Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) with blastp set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Open Gap: 11 and Extension Gap: 1 penalties

30 *Gap x drop-off: 50*

Expect: 10

Word Size: 3

Filter: on

Percent identity may be measured over the length of an entire defined polypeptide sequence,

for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment
5 length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements required for stable mitotic chromosome segregation and maintenance.

10 The term "humanized antibody" refers to antibody molecules in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific
15 hybridization is an indication that two nucleic acid sequences share a high degree of identity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the "washing" step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for
20 annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 µg/ml denatured salmon sperm DNA.

25 Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Generally, such wash temperatures are selected to be about 5°C to 20°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating T_m and
30 conditions for nucleic acid hybridization are well known and can be found in Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; specifically see volume 2, chapter 9.

High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour.

Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, denatured salmon sperm DNA at about 100-200 µg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" and "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

The term "microarray" refers to an arrangement of distinct polynucleotides on a substrate.

The terms "element" and "array element" in a microarray context, refer to hybridizable polynucleotides arranged on the surface of a substrate.

The term "modulate" refers to a change in the activity of hCBP. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of hCBP.

The phrases "nucleic acid" and "nucleic acid sequence" refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

"Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with the second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding

sequence. Generally, operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

"Probe" refers to nucleic acid sequences encoding hCBP, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. "Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the specification, including the tables, figures, and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in the references, for example Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; Ausubel et al., 1987, Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis et al., 1990, PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of

Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to
5 avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing
10 selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or
15 partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

A "recombinant nucleic acid" is a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the
20 artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook, supra. The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to
25 transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be use to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

The term "sample" is used in its broadest sense. A sample suspected of containing nucleic
30 acids encoding hCBP, or fragments thereof, or hCBP itself, may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or

synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide containing the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least about 60% free, preferably about 75% free, and most preferably about 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acids or nucleotides by different amino acids or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

"Transformation" describes a process by which exogenous DNA enters and changes a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed" cells includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95% or at least 98% or greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the

reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 98% or greater sequence identity over a certain defined length of one of the polypeptides.

THE INVENTION

The invention is based on the discovery of a new human calcium binding protein (hCBP), the polynucleotides encoding hCBP, and the use of these compositions for the diagnosis, treatment, or prevention of cancer, reproductive disorders, immune disorders, and developmental disorders.

Nucleic acids encoding the hCBP of the present invention were identified in Incyte Clone 3734805H1 from the coronary artery smooth muscle cDNA library (SMCCNOS01) using a computer search for nucleotide and/or amino acid sequence alignments. A consensus sequence, SEQ ID NO:2, was derived from the following overlapping and/or extended nucleic acid sequences: Incyte Clones 3734805H1 (SMCCNOS01), 3765972F6 (BRSTNOT24), 2260704X21C2 (UTRSNOT02), 2497543T6 (ADRETUT05), and shotgun sequence SBIA06584D1.

In one embodiment, the invention encompasses a polypeptide comprising the amino acid sequence of SEQ ID NO:1, as shown in Figures 1A, 1B, 1C, and 1D. hCBP is 337 amino acids in length and has one potential cAMP- and cGMP-dependent protein kinase phosphorylation site at residue S228; five potential casein kinase II phosphorylation sites at residues S175, T186, T208, S250, and T313; and five potential protein kinase C phosphorylation sites at S10, T35, T186, T224, and S228. PRINTS analysis reveals a putative Huntingtin signature in hCBP within the region from about residues T224 through A243. The gene coding for Huntingtin, a protein associated with Huntington's disease, contains a fragment that is unstable on Huntington's disease chromosomes. As shown in Figures 2A and 2B, hCBP has chemical and structural similarity with MO25 (GI 262934; SEQ ID NO:3), DMO25 (GI 1794137; SEQ ID NO:4), and yeast-like CBP (GI 1255838; SEQ ID NO:5). In particular, hCBP and MO25 share 80% identity, hCBP and DMO25 share 64% identity,

and hCBP and yeast-like CBP share 60% identity. The Huntingtin signature found in hCBP is conserved in all three reference proteins, MO25, DMO25, and yeast-like CBP. In addition, the potential cAMP- and cGMP-dependent protein kinase phosphorylation site at residue S228, the potential casein kinase II phosphorylation sites at residues T186 and S250, and the potential protein kinase C phosphorylation sites at S10, T186, T224, and S228 of hCBP are conserved in MO25, DMO25, and yeast-like CBP. A fragment of SEQ ID NO:2 from about nucleotide 157 to about nucleotide 204 is useful in hybridization or amplification technologies to identify SEQ ID NO:2 and to distinguish between SEQ ID NO:2 and a related sequence. Northern analysis shows the expression of this sequence in various libraries, at least 44% of which are associated with cancer and at least 22% of which are associated with inflammation and the immune response. Of particular note is the expression of hCBP in reproductive tissue.

The invention also encompasses hCBP variants. A preferred hCBP variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the hCBP amino acid sequence, and which contains at least one functional or structural characteristic of hCBP.

The invention also encompasses polynucleotides which encode hCBP. In a particular embodiment, the invention encompasses the polynucleotide sequence consisting of SEQ ID NO:2, which encodes hCBP.

The invention also encompasses a variant of a polynucleotide sequence encoding hCBP. In particular, such a variant polynucleotide sequence will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding hCBP. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising the sequence of SEQ ID NO:2 which has at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence of NO:2. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of hCBP.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding hCBP, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring hCBP, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode hCBP and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring hCBP under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding hCBP or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding hCBP and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode hCBP and hCBP derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding hCBP or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:2 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) *Methods Enzymol.* 152:399-407; Kimmel, A.R. (1987) *Methods Enzymol.* 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Perkin-Elmer), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Perkin-Elmer). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Perkin-Elmer), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995)

Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding hCBP may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Applic. 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060.) Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Perkin-Elmer), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled.

Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode hCBP may be cloned in recombinant DNA molecules that direct expression of hCBP, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express hCBP.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter hCBP-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

In another embodiment, sequences encoding hCBP may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) *Nucleic Acids Symp. Ser.* 7:215-223; and Horn, T. et al. (1980) *Nucleic Acids Symp. Ser.* 7:225-232.) Alternatively, hCBP itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solid-phase techniques. (See, e.g., Roberge, J.Y. et al. (1995) *Science* 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Perkin-Elmer). Additionally, the amino acid sequence of hCBP, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g., Chiez, R.M. and F.Z. Regnier (1990) *Methods Enzymol.* 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY.)

In order to express a biologically active hCBP, the nucleotide sequences encoding hCBP or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding hCBP. Such elements may vary in their strength and specificity. Specific initiation signals

may also be used to achieve more efficient translation of sequences encoding hCBP. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding hCBP and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994)

Results Probl. Cell Differ. 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding hCBP and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding hCBP. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding hCBP. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding hCBP can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPT1 plasmid (Life Technologies). Ligation of sequences encoding hCBP into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of hCBP are needed, e.g. for the production of antibodies, vectors which direct high level expression of hCBP may be used. For example, vectors

containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of hCBP. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such
5 vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra; Bitter, G.A. et al. (1987) *Methods Enzymol.* 153:516-544; and Scorer, C.A. et al. (1994) *Bio/Technology* 12:181-184.)

Plant systems may also be used for expression of hCBP. Transcription of sequences encoding
10 hCBP may be driven viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) *EMBO J.* 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) *EMBO J.* 3:1671-1680; Broglie, R. et al. (1984) *Science* 224:838-843; and Winter, J. et al. (1991) *Results Probl. Cell Differ.* 17:85-105.)
15 These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding hCBP may be ligated into
20 an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses hCBP in host cells. (See, e.g., Logan, J. and T. Shenk (1984) *Proc. Natl. Acad. Sci. USA* 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-
25 based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) *Nat. Genet.*
30 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of hCBP in cell lines is preferred. For example, sequences encoding hCBP can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the

introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk* and *apr* cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), β glucuronidase and its substrate β -glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding hCBP is inserted within a marker gene sequence, transformed cells containing sequences encoding hCBP can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding hCBP under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding hCBP and that express hCBP may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of hCBP using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques

include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on hCBP is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, 5 e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ.)

A wide variety of labels and conjugation techniques are known by those skilled in the art and 10 may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding hCBP include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding hCBP, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, 15 and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic 20 agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding hCBP may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors 25 containing polynucleotides which encode hCBP may be designed to contain signal sequences which direct secretion of hCBP through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, 30 phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct

modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding hCBP may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric hCBP protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of hCBP activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the hCBP encoding sequence and the heterologous protein sequence, so that hCBP may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, *supra*, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled hCBP may be achieved *in vitro* using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, ³⁵S-methionine.

Fragments of hCBP may be produced not only by recombinant means, but also by direct peptide synthesis using solid-phase techniques. (See, e.g., Creighton, *supra*, pp. 55-60.) Protein synthesis may be performed by manual techniques or by automation. Automated synthesis may be achieved, for example, using the ABI 431A peptide synthesizer (Perkin-Elmer). Various fragments of hCBP may be synthesized separately and then combined to produce the full length molecule.

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of hCBP, MO25 (GI 262934), DMO25 (GI 1794137), and yeast-like CBP (GI 1255838). In addition, the expression of hCBP is closely associated with cancer, reproductive tissues, inflammation and the immune response. Therefore, hCBP appears to play a role in cancer, reproductive disorders,

immune disorders, and developmental disorders. In the treatment of disorders associated with increased hCBP expression or activity, it is desirable to decrease the expression or activity of hCBP. In the treatment of disorders associated with decreased hCBP expression or activity, it is desirable to increase the expression or activity of hCBP.

- 5 Therefore, in one embodiment, hCBP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of hCBP. Examples of such disorders include, but are not limited to, a cancer, such as adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, a cancer of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia,
- 10 gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; a reproductive disorder such as disorders of prolactin production, infertility, including tubal disease, ovulatory defects, and endometriosis, disruptions of the estrous cycle, disruptions of the menstrual cycle, polycystic ovary syndrome, ovarian hyperstimulation syndrome, endometrial and ovarian tumors, uterine fibroids,
- 15 autoimmune disorders, ectopic pregnancies, and teratogenesis, cancer of the breast, fibrocystic breast disease, and galactorrhea, disruptions of spermatogenesis, abnormal sperm physiology, cancer of the testis, cancer of the prostate, benign prostatic hyperplasia, prostatitis, Peyronie's disease, impotence, carcinoma of the male breast, and gynecomastia; a developmental disorder, such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular
- 20 dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucoepithelial dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Sydenham's chorea and cerebral palsy, spina bifida, anencephaly,
- 25 craniorachischisis, congenital glaucoma, cataract, and sensorineural hearing loss; and an immune disorder, such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-
- 30 candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's

syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma.

5 In another embodiment, a vector capable of expressing hCBP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of hCBP including, but not limited to, those described above.

In a further embodiment, a pharmaceutical composition comprising a substantially purified hCBP in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat
10 or prevent a disorder associated with decreased expression or activity of hCBP including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of hCBP may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of hCBP including, but not limited to, those listed above.

15 In a further embodiment, an antagonist of hCBP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of hCBP. Examples of such disorders include, but are not limited to, those cancer, reproductive disorders, immune disorders, and developmental disorders described above. In one aspect, an antibody which specifically binds hCBP may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a
20 pharmaceutical agent to cells or tissues which express hCBP.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding hCBP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of hCBP including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary
25 sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic
30 efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of hCBP may be produced using methods which are generally known in the art. In particular, purified hCBP may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind hCBP. Antibodies to hCBP may also be generated using methods that are well known in the art. Such antibodies may include, but are not

limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans,
5 and others may be immunized by injection with hCBP or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans,
10 BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to hCBP have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein and contain the
15 entire amino acid sequence of a small, naturally occurring molecule. Short stretches of hCBP amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to hCBP may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not
20 limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; and Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the
25 splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce hCBP-specific single
30 chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as

disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for hCBP may also be generated. For example, such fragments include, but are not limited to, F(ab')₂ fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')₂ fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between hCBP and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering hCBP epitopes is generally used, but a competitive binding assay may also be employed (Pound, supra).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for hCBP. Affinity is expressed as an association constant, K_a, which is defined as the molar concentration of hCBP-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple hCBP epitopes, represents the average affinity, or avidity, of the antibodies for hCBP. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular hCBP epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10⁹ to 10¹² L/mole are preferred for use in immunoassays in which the hCBP-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10⁶ to 10⁷ L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of hCBP, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington, DC; Liddell, J.E. and Cryer, A. (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of hCBP-

antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al. supra.)

In another embodiment of the invention, the polynucleotides encoding hCBP, or any fragment
5 or complement thereof, may be used for therapeutic purposes. In one aspect, the complement of the polynucleotide encoding hCBP may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding hCBP. Thus, complementary molecules or fragments may be used to modulate hCBP activity, or to achieve regulation of gene function. Such technology is now well
10 known in the art, and sense or antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding hCBP.

Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. Methods which are well known to those skilled in the art can be used
15 to construct vectors to express nucleic acid sequences complementary to the polynucleotides encoding hCBP. (See, e.g., Sambrook, supra; Ausubel, 1995, supra.)

Genes encoding hCBP can be turned off by transforming a cell or tissue with expression vectors which express high levels of a polynucleotide, or fragment thereof, encoding hCBP. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in
20 the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector, and may last even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing
25 complementary sequences or antisense molecules (DNA, RNA, or PNA) to the control, 5', or regulatory regions of the gene encoding hCBP. Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may be employed. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the
30 binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze
5 endonucleolytic cleavage of sequences encoding hCBP.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for
10 secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques
15 for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA sequences encoding hCBP. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines,
20 cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be
25 extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Many methods for introducing vectors into cells or tissues are available and equally suitable for use *in vivo*, *in vitro*, and *ex vivo*. For *ex vivo* therapy, vectors may be introduced into stem cells
30 taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of

such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and monkeys.

An additional embodiment of the invention relates to the administration of a pharmaceutical or sterile composition, in conjunction with a pharmaceutically acceptable carrier, for any of the
5 therapeutic effects discussed above. Such pharmaceutical compositions may consist of hCBP, antibodies to hCBP, and mimetics, agonists, antagonists, or inhibitors of hCBP. The compositions may be administered alone or in combination with at least one other agent, such as a stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered
10 to a patient alone, or in combination with other agents, drugs, or hormones.

The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

15 In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA).

20 Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combining active
25 compounds with solid excipient and processing the resultant mixture of granules (optionally, after grinding) to obtain tablets or dragee cores. Suitable auxiliaries can be added, if desired. Suitable excipients include carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, and sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums, including arabic and
30 tragacanth; and proteins, such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, and alginic acid or a salt thereof, such as sodium alginate.

Dragee cores may be used in conjunction with suitable coatings, such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene

glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, i.e., dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with fillers or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils, such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate, triglycerides, or liposomes. Non-lipid polycationic amino polymers may also be used for delivery. Optionally, the suspension may also contain suitable stabilizers or agents to increase the solubility of the compounds and allow for the preparation of highly concentrated solutions.

For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention may be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

The pharmaceutical composition may be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, and succinic acids. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preparation may be a lyophilized powder which may contain any or all of the following: 1 mM to 50 mM histidine, 0.1% to 2% sucrose, and 2% to 7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of hCBP, such labeling would include amount, frequency, and method of administration.

Pharmaceutical compositions suitable for use in the invention include compositions wherein

the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, or pigs.

- 5 An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

- A therapeutically effective dose refers to that amount of active ingredient, for example hCBP or fragments thereof, antibodies of hCBP, and agonists, antagonists or inhibitors of hCBP, which
10 ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD_{50}/ED_{50} ratio. Pharmaceutical compositions
15 which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

- 20 The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and
25 response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

- Normal dosage amounts may vary from about 0.1 μg to 100,000 μg , up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and
30 methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind hCBP may be used for the diagnosis of disorders characterized by expression of hCBP, or in assays to monitor patients being treated with hCBP or agonists, antagonists, or inhibitors of hCBP. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for hCBP include methods which utilize the antibody and a label to detect hCBP in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring hCBP, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of hCBP expression. Normal or standard values for hCBP expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibody to hCBP under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of hCBP expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding hCBP may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of hCBP may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of hCBP, and to monitor regulation of hCBP levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding hCBP or closely related molecules may be used to identify nucleic acid sequences which encode hCBP. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding hCBP, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the hCBP encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:2 or from genomic sequences including promoters, enhancers, and introns of the hCBP gene.

Means for producing specific hybridization probes for DNAs encoding hCBP include the cloning of polynucleotide sequences encoding hCBP or hCBP derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA

- 5 polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S , or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

- Polynucleotide sequences encoding hCBP may be used for the diagnosis of disorders associated with expression of hCBP. Examples of such disorders include, but are not limited to, a
- 10 cancer, such as adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, a cancer of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; a reproductive disorder such as disorders of prolactin production, infertility, including tubal disease,
- 15 ovulatory defects, and endometriosis, disruptions of the estrous cycle, disruptions of the menstrual cycle, polycystic ovary syndrome, ovarian hyperstimulation syndrome, endometrial and ovarian tumors, uterine fibroids, autoimmune disorders, ectopic pregnancies, and teratogenesis, cancer of the breast, fibrocystic breast disease, and galactorrhea, disruptions of spermatogenesis, abnormal sperm physiology, cancer of the testis, cancer of the prostate, benign prostatic hyperplasia, prostatitis,
- 20 Peyronie's disease, impotence, carcinoma of the male breast, and gynecomastia; a developmental disorder, such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucoepithelial dysplasia, hereditary keratodermas, hereditary
- 25 neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Sydenham's chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, and sensorineural hearing loss; and an immune disorder, such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma,
- 30 atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's

thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma. The polynucleotide sequences encoding hCBP may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered hCBP expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding hCBP may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding hCBP may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding hCBP in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of hCBP, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding hCBP, under conditions suitable for hybridization or amplification.

Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding hCBP may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding hCBP, or a fragment of a polynucleotide complementary to the polynucleotide encoding hCBP, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

Methods which may also be used to quantify the expression of hCBP include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) *J. Immunol. Methods* 159:235-244; Duplaa, C. et al. (1993) *Anal. Biochem.* 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as targets in a microarray. The microarray can be used to monitor the expression level of large numbers of genes simultaneously and to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, and to develop and monitor the activities of therapeutic agents.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) *Proc. Natl. Acad. Sci. USA* 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) *Proc. Natl. Acad. Sci. USA* 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.)

In another embodiment of the invention, nucleic acid sequences encoding hCBP may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to

artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.)

5 Fluorescent in situ hybridization (FISH) may be correlated with other physical chromosome mapping techniques and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding hCBP on a physical chromosomal map and a specific disorder, or a
10 predisposition to a specific disorder, may help define the region of DNA associated with that disorder. The nucleotide sequences of the invention may be used to detect differences in gene sequences among normal, carrier, and affected individuals.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps.

15 Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic
20 linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

25 In another embodiment of the invention, hCBP, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between hCBP and the agent being tested may be measured.

30 Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with hCBP, or fragments thereof, and washed. Bound hCBP is then detected by methods well known in the art. Purified hCBP can also

be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding hCBP specifically compete with a test compound for binding hCBP. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with hCBP.

In additional embodiments, the nucleotide sequences which encode hCBP may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, in particular U.S. Ser. No. 09/190,965, are hereby expressly incorporated by reference.

EXAMPLES

I. Construction of cDNA Libraries

The SMCCNOS01 subtracted coronary artery smooth muscle cell library was constructed using 7.56×10^6 clones from the SMCCNOT02 library and was subjected to two rounds of subtraction hybridization for 48 hours with 6.12×10^6 clones from SMCCNOT01. The starting library for subtraction was constructed using RNA isolated from coronary artery smooth muscle cells removed from a 3-year-old Caucasian male. The cells were treated with TNF α and IL-1 β , 10ng/ml each, for 20 hours. The hybridization probe for subtraction was derived from a similarly constructed library from RNA isolated from untreated coronary artery smooth muscle cells from the same donor. Subtractive hybridization conditions were based on the methodologies of Swaroop, A. et al. (1991, Nucleic Acids Res. 19:1954) and Bonaldo, M.F. et al. (1996, Genome Res. 6:791-806).

For both cDNA libraries SMCCNOT01 and SMCCNOT02, the frozen coronary artery smooth muscle cells (50-100 mg) were homogenized in GTC homogenization buffer (4.0M guanidine thiocyanate, 0.1M Tris-HCl pH 7.5, 1% 2-mercaptoethanol). Two volumes of binding buffer (0.4M LiCl, 0.1M Tris-HCl pH 7.5, 0.02M EDTA) were added and the resulting mixture vortexed at 13,000 rpm. The supernatant was removed and combined with Oligo(dT)₂₅ bound MPG Streptavidin particles. After rotation at room temperature, the mRNA-Oligo(dT)₂₅-Streptavidin particles were

separated from the supernatant, washed twice with hybridization buffer I (0.15M NaCl, 0.01M Tris-HCl pH8.0, 1mM EDTA, 0.1% lauryl sarcosinate) and washed twice with hybridization buffer II (0.15M NaCl, 0.01M Tris-HCl pH8.0, 1 mM EDTA) using magnetic separation at each step to remove the supernatant from the particles. Bound mRNA was eluted from the MPG Streptavidin
5 particles with release solution and heated to 65°C. The supernatant containing eluted mRNA was magnetically separated from the Streptavidin particles and used to construct the cDNA libraries.

Poly(A+) RNA was used for cDNA synthesis and construction of the cDNA library according to the recommended protocols in the SUPERScript plasmid system (Life Technologies). The cDNAs were fractionated on a SEPHAROSE CL4B column (Amersham Pharmacia Biotech), and
10 those cDNAs exceeding 400 bp were ligated into pINCY (Incyte Pharmaceuticals, Palo Alto CA). Recombinant plasmids were transformed into DH5α competent cells or ElectroMAX® cells (Life Technologies).

II. Isolation of cDNA Clones

Plasmids were recovered from host cells by in vivo excision using the UNIZAP vector system
15 (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or
20 without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically
25 using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

cDNA sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Perkin-Elmer) thermal cycler or the PTC-200
30 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer).

Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides

were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (Perkin-Elmer) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, supra, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example V.

The polynucleotide sequences derived from cDNA sequencing were assembled and analyzed using a combination of software programs which utilize algorithms well known to those skilled in the art. Table 1 summarizes the tools, programs, and algorithms used and provides applicable descriptions, references, and threshold parameters. The first column of Table 1 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score, the greater the homology between two sequences). Sequences were analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments were generated using the default parameters specified by the clustal algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

The polynucleotide sequences were validated by removing vector, linker, and polyA sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The sequences were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM, and PFAM to acquire annotation using programs based on BLAST, FASTA, and BLIMPS. The sequences were assembled into full length polynucleotide sequences using programs based on Phred, Phrap, and Consed, and were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length amino acid sequences, and these full length sequences were subsequently analyzed by querying against databases such as the GenBank databases (described above), SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and Hidden Markov Model (HMM)-based protein family databases such as PFAM. HMM is a probabilistic approach which analyzes consensus primary structures of gene families. (See, e.g., Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365.)

The programs described above for the assembly and analysis of full length polynucleotide and amino acid sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:2. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies were described in The Invention section above.

5 IV. Northern Analysis

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, supra, ch. 7; Ausubel, 1995, supra, ch. 4 and 16.)

10 Analogous computer techniques applying BLAST were used to search for identical or related molecules in nucleotide databases such as GenBank or LIFESEQ (Incyte Pharmaceuticals). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$15 \quad \frac{\% \text{ sequence identity} \times \% \text{ maximum BLAST score}}{100}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1% to 2% error, and, with a product score of 70, the match will be exact. Similar molecules are
20 usually identified by selecting those which show product scores between 15 and 40, although lower scores may identify related molecules.

The results of northern analyses are reported as a percentage distribution of libraries in which the transcript encoding hCBP occurred. Analysis involved the categorization of cDNA libraries by organ/tissue and disease. The organ/tissue categories included cardiovascular, dermatologic,
25 developmental, endocrine, gastrointestinal, hematopoietic/immune, musculoskeletal, nervous, reproductive, and urologic. The disease/condition categories included cancer, inflammation, trauma, cell proliferation, neurological, and pooled. For each category, the number of libraries expressing the sequence of interest was counted and divided by the total number of libraries across all categories. Percentage values of tissue-specific and disease- or condition-specific expression are reported in the
30 description of the invention.

V. Extension of hCBP Encoding Polynucleotides

The full length nucleic acid sequences of SEQ ID NO:2 were produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other

primer, to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would
5 result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction
10 mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+} , $(NH_4)_2SO_4$, and β -mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the
15 alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 μ l PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE
20 and 0.5 μ l of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 μ l to 10 μ l aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose mini-gel to determine which reactions were successful in extending
25 the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For
30 shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent *E. coli* cells. Transformed cells were selected on antibiotic-containing media, individual colonies were picked and cultured overnight at 37°C in 384-

well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethylsulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer).

In like manner, the nucleotide sequences of SEQ ID NO:2 are used to obtain 5' regulatory sequences using the procedure above, oligonucleotides designed for such extension, and an appropriate genomic library.

VI. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:2 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μ Ci of [γ -³²P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10⁷ counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

VII. Microarrays

A chemical coupling procedure and an ink jet device can be used to synthesize array elements on the surface of a substrate. (See, e.g., Baldeschweiler, supra.) An array analogous to a dot

or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced by hand or using available methods and machines and contain any appropriate number of elements. After hybridization, nonhybridized probes are removed and a scanner used to determine the levels and patterns of fluorescence. The degree of complementarity and the relative abundance of each probe which hybridizes to an element on the microarray may be assessed through analysis of the scanned images.

Full-length cDNAs, Expressed Sequence Tags (ESTs), or fragments thereof may comprise the elements of the microarray. Fragments suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). Full-length cDNAs, ESTs, or fragments thereof corresponding to one of the nucleotide sequences of the present invention, or selected at random from a cDNA library relevant to the present invention, are arranged on an appropriate substrate, e.g., a glass slide. The cDNA is fixed to the slide using, e.g., UV cross-linking followed by thermal and chemical treatments and subsequent drying. (See, e.g., Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645.) Fluorescent probes are prepared and used for hybridization to the elements on the substrate. The substrate is analyzed by procedures described above.

VIII. Complementary Polynucleotides

Sequences complementary to the hCBP-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring hCBP. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of hCBP. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the hCBP-encoding transcript.

IX. Expression of hCBP

Expression and purification of hCBP is achieved using bacterial or virus-based expression systems. For expression of hCBP in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express hCBP upon induction with isopropyl beta-D-

thiogalactopyranoside (IPTG). Expression of hCBP in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding hCBP by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, hCBP is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from hCBP at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch. 10 and 16). Purified hCBP obtained by these methods can be used directly in the following activity assay.

X. Demonstration of hCBP Activity

The activity assay for human calcium-binding proteins is based upon the ability of hCBPs to down-regulate mitosis. hCBP can be expressed by transforming a mammalian cell line such as COS7, HeLa or CHO with a eukaryotic expression vector encoding hCBP. Eukaryotic expression vectors are commercially available, and the techniques to introduce them into cells are well known to those skilled in the art. The cells are incubated for 48-72 hours after transformation under conditions appropriate for the cell line to allow expression of hCBP. Then, phase microscopy is used to compare the mitotic index of transformed versus control cells. The decrease in the mitotic index is proportional to the hCBP activity.

XI. Functional Assays

hCBP function is assessed by expressing the sequences encoding hCBP at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression

vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT (Life Technologies) and pCR3.1 (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10 μ g of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 μ g of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of hCBP on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding hCBP and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding hCBP and other genes of interest can be analyzed by northern analysis or microarray techniques.

XII. Production of hCBP Specific Antibodies

hCBP substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) *Methods Enzymol.* 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the hCBP amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for

selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Perkin-Elmer) using fmoc-chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for anti-peptide and anti-hCBP activity by, for example, binding the peptide or hCBP to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

10 XIII. Purification of Naturally Occurring hCBP Using Specific Antibodies

Naturally occurring or recombinant hCBP is substantially purified by immunoaffinity chromatography using antibodies specific for hCBP. An immunoaffinity column is constructed by covalently coupling anti-hCBP antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is
15 blocked and washed according to the manufacturer's instructions.

Media containing hCBP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of hCBP (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/hCBP binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as
20 urea or thiocyanate ion), and hCBP is collected.

XIV. Identification of Molecules Which Interact with hCBP

hCBP, or biologically active fragments thereof, are labeled with ¹²⁵I Bolton-Hunter reagent. (See, e.g., Bolton A.E. and W.M. Hunter (1973) Biochem. J. 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled hCBP, washed, and
25 any wells with labeled hCBP complex are assayed. Data obtained using different concentrations of hCBP are used to calculate values for the number, affinity, and association of hCBP with the candidate molecules.

Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention.
30 Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

Table 1

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25: 3389-3402.	ESTs: Probability value= 1.0E-8 or less Full Length sequences: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises at least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad. Sci. 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183: 63-98; and Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta E value=1.06E-6 Assembled ESTs: fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less Full Length sequences: fastx score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S and J.G. Henikoff, Nucl. Acid Res. 19:6565-72, 1991. J.G. Henikoff and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37: 417-424.	Score=1000 or greater; Ratio of Score/Strength = 0.75 or larger; and, if applicable, Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol. 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322.	Score=10-50 bits for PFAM hits, depending on individual protein families

Table 1 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25: 217-221.	Normalized quality score > GCG-specified "HIGH" value for that particular Prosite motif. Generally, score = 1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M. S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score = 120 or greater; Match length = 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Clavette, J.M. and S. Audic (1997) CABIOS 12: 431-439.	Score = 3.5 or greater
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch et al. <u>supra</u> ; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

What is claimed is:

1. A substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1 and fragments thereof.
5
2. A substantially purified variant having at least 90% amino acid sequence identity to the amino acid sequence of claim 1.
3. An isolated and purified polynucleotide encoding the polypeptide of claim 1.
10
4. An isolated and purified polynucleotide variant having at least 70% polynucleotide sequence identity to the polynucleotide of claim 3.
5. An isolated and purified polynucleotide which hybridizes under stringent conditions
15 to the polynucleotide of claim 3.
6. An isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide of claim 3.
- 20 7. A method for detecting a polynucleotide, the method comprising the steps of:
 - (a) hybridizing the polynucleotide of claim 6 to at least one nucleic acid in a sample, thereby forming a hybridization complex; and
 - (b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of the polynucleotide in the sample.
25
8. The method of claim 7 further comprising amplifying the polynucleotide prior to hybridization.
9. An isolated and purified polynucleotide comprising a polynucleotide sequence
30 selected from the group consisting of SEQ ID NO:2 and fragments thereof.
10. An isolated and purified polynucleotide variant having at least 70% polynucleotide sequence identity to the polynucleotide of claim 9.

11. An isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide of claim 9.

5

12. An expression vector comprising at least a fragment of the polynucleotide of claim 3.

13. A host cell comprising the expression vector of claim 12.

14. A method for producing a polypeptide, the method comprising the steps of:

10

a) culturing the host cell of claim 13 under conditions suitable for the expression of the polypeptide; and

b) recovering the polypeptide from the host cell culture.

15. A pharmaceutical composition comprising the polypeptide of claim 1 in conjunction with a suitable pharmaceutical carrier.

15

16. A purified antibody which specifically binds to the polypeptide of claim 1.

17. A purified agonist of the polypeptide of claim 1.

20

18. A purified antagonist of the polypeptide of claim 1.

19. A method for treating or preventing a disorder associated with decreased expression or activity of hCBP, the method comprising administering to a subject in need of such treatment an effective amount of the pharmaceutical composition of claim 15.

25

20. A method for treating or preventing a disorder associated with increased expression or activity of hCBP, the method comprising administering to a subject in need of such treatment an effective amount of the antagonist of claim 18.

1/6

```

5' CAG AAA CAG AAC TGC CTG TGA CAG ATT AAG AGA CAA GCA AGG CTT GGA ATC TGA
    9      18      27      36      45      54
GAG CAA GCA AAG AGA GTG GAA ATT TAC AGC TGC CTT ATC ATT CCA TAT TGG AAG
    63      72      81      90      99      108
AAG AGA TTT CTA CAC ATG AAA AAA ATG CCT TTG TTT AGT AAA TCA CAC AAA AAT
    117     126     135     144     153     162
CCA GCA GAA ATT GTG AAA ATC CTG AAA GAC AAT TTG GCC ATT TTG GAA AAG CAA
    171     180     189     198     207     216
GAC AAA AAG ACA GAC AAG GCT TCA GAA GAA GTG TCT AAA TCA CTG CAA GCA ATG
    225     234     243     252     261     270
AAA GAA ATT CTG TGT GGT ACA AAC GAG AAA GAA CCC CCG ACA GAA GCA GTG GCT
    279     288     297     306     315     324
CAG CTA GCA CAA GAA CTC TAC AGC AGT GGC CTG CTG CTG GTG ACA CTG ATA GCT GAC
    333     342     351     360     369     378
Q  L  A  Q  E  L  Y  S  S  G  L  L  L  V  T  L  I  A  D

```

FIGURE 1A

387	CTG CAG CTG ATA GAC TTT GAG GGA AAA AAA GAT GTG ACC CAG ATA TTT AAC AAC	432
	L Q L I D F E G K K K D V T Q I F N N	
441	ATC TTG AGA AGA CAG ATA GGC ACT CGG AGT CCT ACT GTG GAG TAT ATT AGT GCT	486
	I L R R Q I G T R S P T T V E Y I S A	
495	CAT CCT CAT ATC CTG TTT ATG CTC CTC AAA GGA TAT GAA GCC CCA CAG ATT GCC	540
	H P H I L F M L L L K G Y E A P Q I A	
549	TTA CGT TGT GGG ATT ATG CTG AGA GAA TGT ATT CGA CAT GAA CCA CTT GCC AAA	594
	L R C G I M L R E C I R H E P L A K	
603	ATC ATC CTC TTT TCT AAT CAA TTC AGA GAT TTC TTT AAG TAC GTG GAG TTG TCA	648
	I I L F S N Q F R D F F F K Y V E L S	
657	ACA TTT GAT ATT GCT TCA GAT GCC TTT GCT ACT TTC AAG GAT TTA CTA ACC AGA	702
	T F D I A S D A F A T F F K D L L T R	
711	CAT AAA GTG TTG GTA GCA GAC TTC TTA GAA CAA AAT TAC GAC ACT ATT TTT GAA	756
	H K V L V A D F L L E Q N Y D T I F E	

FIGURE 1B

765	774	783	792	801	810
GAC TAT GAG AAA TTG CTT CAG TCT GAG AAT TAT GTT ACT AAG AGA CAG TCT TTA					
D Y E K L L L Q S E N Y V T K R Q S L					
819	828	837	846	855	864
AAG CTG CTA GGG GAG CTG ATC CTG GAC CGT CAC AAC TTT GCC ATC ATG ACA AAG					
K L L G E L I L D R H N F A I M T K					
873	882	891	900	909	918
TAT ATC AGC AAG CCG GAG AAC CTG AAA CTC ATG ATG AAC CTC CTG CGG GAT AAA					
Y I S K P E N L K L L M M N L L R D K					
927	936	945	954	963	972
AGT CCC AAC ATC CAG TTT GAA GCC TTT CAT GTT TTT AAG GTG TTT GTG GCC AGT					
S P N I Q F E A F H V F K V F V A S					
981	990	999	1008	1017	1026
CCT CAC AAA ACA CAG CCT ATT GTG GAG ATC CTG TTA AAA AAT CAG CCC AAA CTC					
P H K T Q P I V E I L L K N Q P K L					
1035	1044	1053	1062	1071	1080
ATT GAG TTT CTG AGC AGC TTC CAA AAA GAA AGG ACG GAT GAT GAG CAG TTC GCT					
I E F L S S F Q K E E R T D E Q F A					
1089	1098	1107	1116	1125	1134
GAC GAG AAG AAC TAC TTG ATT AAA CAG ATC CGA GAC TTG AAG AAA ACG GCC CCT					
D E K N Y L I K Q I R D L K K T A P					

FIGURE 1C

1143 1152 1161 1170 1179 1188
TGA AGA GCT CCC CGG CCC CTG TCA CAG TCA GTC GTC TCA TTT GTC CAG TTT GTA

1197 1206 1215 1224 1233 1242
CAG TGT GTC ATT TCA GAA AGT CAT CAT TCT TGG GAA GAC TTT GGA GGT GCC TAT

1251 1260 1269 1278 1287 1296
TTT TTC TGC TGT AAT TGT TCT GGG TAG ATG GAG TAT AAA CAT TTG AAT GGA AAA

1305 1314 1323 1332 1341
AAA TTA ACC TAG AAT AAT ATA TTC ATT TTA GTC AAA AAA AAA AAA AAA 3'

FIGURE 1D

[illegible]

FIGURE 2A

183	A	F	A	T	F	K	D	L	L	T	R	H	K	V	L	V	A	D	F	F	L	E	Q	N	Y	D	T	I	F	E	D	-	Y	E	K	L	L	Q	S	E	N	hCBP
184	A	F	A	T	F	K	D	L	L	T	R	H	K	L	L	S	A	E	F	F	L	E	Q	H	Y	D	R	F	F	S	E	-	Y	E	K	L	L	H	S	E	N	g262934
183	A	F	S	T	F	K	E	L	L	T	R	H	K	L	L	C	A	E	F	L	D	A	N	Y	D	K	F	F	S	Q	H	-	Y	Q	R	L	L	N	S	E	N	g1794137
197	A	F	S	T	F	K	E	L	L	T	R	H	K	A	I	A	E	F	L	D	S	N	Y	D	T	F	F	F	A	Q	-	Y	Q	N	L	L	N	S	K	N	g1255838	
222	Y	V	T	K	R	Q	S	L	K	L	L	G	E	L	I	L	D	R	H	N	F	A	I	M	T	K	Y	I	S	K	P	E	N	L	K	L	M	M	N	L	hCBP	
223	Y	V	T	K	R	Q	S	L	K	L	L	G	E	L	L	L	D	R	H	N	F	T	I	M	T	K	Y	I	S	K	P	E	N	L	K	L	M	M	N	L	g262934	
223	Y	V	T	R	R	Q	S	L	K	L	L	G	E	L	L	L	D	R	H	N	F	T	V	M	T	R	Y	I	S	E	P	E	N	L	K	L	M	M	N	M	g1794137	
236	Y	V	T	R	R	Q	S	L	K	L	L	G	E	L	L	L	D	R	H	N	F	N	T	M	T	K	Y	I	S	N	P	D	N	L	R	L	M	M	E	L	g1255838	
262	L	R	D	K	S	P	N	I	Q	F	E	A	F	H	V	F	K	V	F	V	A	S	P	H	K	T	Q	P	I	V	E	I	L	L	K	N	Q	P	K	L	hCBP	
263	L	R	D	K	S	R	N	I	Q	F	E	A	F	H	V	F	K	V	F	V	A	N	P	N	K	T	Q	P	I	L	D	I	L	L	K	N	Q	T	K	L	g262934	
263	L	K	E	K	S	R	N	I	Q	F	E	A	F	H	V	F	K	V	F	V	A	N	P	N	K	P	K	P	I	L	D	I	L	L	R	N	Q	T	K	L	g1794137	
276	L	R	D	K	S	R	N	I	Q	Y	E	A	F	H	V	F	K	V	F	V	A	N	P	N	K	P	K	P	I	S	D	I	L	N	R	N	R	E	K	L	g1255838	
302	I	E	F	L	S	S	F	Q	K	E	R	T	D	D	E	Q	F	A	D	E	K	N	Y	L	I	K	Q	I	R	D	L	K	K	T	A	-	-	-	-	-	hCBP	
303	I	E	F	L	S	K	F	Q	N	D	R	T	E	D	E	Q	F	N	D	E	K	T	Y	L	V	K	Q	I	R	N	L	K	R	A	A	Q	Q	E	-	-	g262934	
303	V	D	F	L	T	N	F	H	T	D	R	S	E	D	E	Q	F	N	D	E	K	A	Y	L	I	K	Q	I	K	E	L	K	P	L	P	-	-	-	-	g1794137		
316	V	E	F	L	S	E	F	H	N	D	R	T	D	D	E	Q	F	N	D	E	K	A	Y	L	I	K	Q	I	Q	E	M	K	S	S	P	K	E	A	K	g1255838		
337	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	hCBP			
341	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	g262934			
339	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	g1794137			
356	P	K	S	K	E	D	E	N	Q	E	P	A	G	P	S	E	G	P	S	T	S	Q	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	g1255838		

FIGURE 2B

SEQUENCE LISTING

<110> INCYTE PHARMACEUTICALS, INC.

TANG, Y. Tom

GUEGLER, Karl J.

CORLEY, Neil C.

GORGONE, Gina A.

<120> CALCIUM BINDING PROTEIN

<130> PF-0635 PCT

<140> To Be Assigned

<141> Herewith

<150> 09/190,965

<151> 1998-11-13

<160> 5

<170> PERL Program

<210> 1

<211> 337

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3734805CD1

<400> 1

Met	Lys	Lys	Met	Pro	Leu	Phe	Ser	Lys	Ser	His	Lys	Asn	Pro	Ala
1				5					10					15
Glu	Ile	Val	Lys	Ile	Leu	Lys	Asp	Asn	Leu	Ala	Ile	Leu	Glu	Lys
			20						25					30
Gln	Asp	Lys	Lys	Thr	Asp	Lys	Ala	Ser	Glu	Glu	Val	Ser	Lys	Ser
			35						40					45
Leu	Gln	Ala	Met	Lys	Glu	Ile	Leu	Cys	Gly	Thr	Asn	Glu	Lys	Glu
			50						55					60
Pro	Pro	Thr	Glu	Ala	Val	Ala	Gln	Leu	Ala	Gln	Glu	Leu	Tyr	Ser
			65						70					75
Ser	Gly	Leu	Leu	Val	Thr	Leu	Ile	Ala	Asp	Leu	Gln	Leu	Ile	Asp
			80						85					90
Phe	Glu	Gly	Lys	Lys	Asp	Val	Thr	Gln	Ile	Phe	Asn	Asn	Ile	Leu
			95						100					105
Arg	Arg	Gln	Ile	Gly	Thr	Arg	Ser	Pro	Thr	Val	Glu	Tyr	Ile	Ser
			110						115					120
Ala	His	Pro	His	Ile	Leu	Phe	Met	Leu	Leu	Lys	Gly	Tyr	Glu	Ala
			125						130					135
Pro	Gln	Ile	Ala	Leu	Arg	Cys	Gly	Ile	Met	Leu	Arg	Glu	Cys	Ile
			140						145					150
Arg	His	Glu	Pro	Leu	Ala	Lys	Ile	Ile	Leu	Phe	Ser	Asn	Gln	Phe
			155						160					165

Arg	Asp	Phe	Phe	Lys	Tyr	Val	Glu	Leu	Ser	Thr	Phe	Asp	Ile	Ala	
				170					175					180	
Ser	Asp	Ala	Phe	Ala	Thr	Phe	Lys	Asp	Leu	Leu	Thr	Arg	His	Lys	
				185					190					195	
Val	Leu	Val	Ala	Asp	Phe	Leu	Glu	Gln	Asn	Tyr	Asp	Thr	Ile	Phe	
				200					205					210	
Glu	Asp	Tyr	Glu	Lys	Leu	Leu	Gln	Ser	Glu	Asn	Tyr	Val	Thr	Lys	
				215					220					225	
Arg	Gln	Ser	Leu	Lys	Leu	Leu	Gly	Glu	Leu	Ile	Leu	Asp	Arg	His	
				230					235					240	
Asn	Phe	Ala	Ile	Met	Thr	Lys	Tyr	Ile	Ser	Lys	Pro	Glu	Asn	Leu	
				245					250					255	
Lys	Leu	Met	Met	Asn	Leu	Leu	Arg	Asp	Lys	Ser	Pro	Asn	Ile	Gln	
				260					265					270	
Phe	Glu	Ala	Phe	His	Val	Phe	Lys	Val	Phe	Val	Ala	Ser	Pro	His	
				275					280					285	
Lys	Thr	Gln	Pro	Ile	Val	Glu	Ile	Leu	Leu	Lys	Asn	Gln	Pro	Lys	
				290					295					300	
Leu	Ile	Glu	Phe	Leu	Ser	Ser	Phe	Gln	Lys	Glu	Arg	Thr	Asp	Asp	
				305					310					315	
Glu	Gln	Phe	Ala	Asp	Glu	Lys	Asn	Tyr	Leu	Ile	Lys	Gln	Ile	Arg	
				320					325					330	
Asp	Leu	Lys	Lys	Thr	Ala	Pro									
				335											

<210> 2

<211> 1344

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3734805CB1

<400> 2

cagaaacaga	actgcctgtg	acagattaag	agacaagcaa	ggcttggaat	ctgagagcaa	60
gcaaagagag	tggaaattta	cagctgcctt	atcattccat	attggaagaa	gagatttcta	120
cacatgaaaa	aaatgccttt	gttttagtaa	tcacacaaaa	atccagcaga	aattgtgaaa	180
atcctgaaag	acaatttggc	cattttggaa	aagcaagaca	aaaagacaga	caaggcttca	240
gaagaagtgt	ctaaatcact	gcaagcaatg	aaagaaattc	tgtgtggtac	aaacgagaaa	300
gaacccccga	cagaagcagt	ggctcagcta	gcacaagaac	tctacagcag	tggcctgctg	360
gtgacactga	tagctgacct	gcagctgata	gactttgagg	gaaaaaaaga	tgtgacctag	420
atatttaaca	acatcttgag	aagacagata	ggcactcgga	gtcctactgt	ggagtattat	480
agtgtctatc	ctcatatcct	gtttatgctc	ctcaaaggat	atgaagcccc	acagattgcc	540
ttacgttggtg	ggattatgct	gagagaatgt	attcgacatg	aaccacttgc	caaaatcatc	600
ctcttttcta	atcaattcag	agattttctt	aagtacgtgg	agttgtcaac	atttgatatt	660
gcttcagatg	cctttgctac	tttcaaggat	ttactaacca	gacataaagt	gttggttagca	720
gacttccttag	aacaaaatta	cgacactatt	tttgaagact	atgagaaatt	gcttcagtct	780
gagaattatg	ttactaagag	acagtcttta	aagctgctag	gggagctgat	cctggaccgt	840
cacaactttg	ccatcatgac	aaagtatatc	agcaagccgg	agaacctgaa	actcatgatg	900
aacctccttc	gggataaaa	tcccaacatc	cagtttgaa	cctttcatgt	ttttaagggtg	960
tttgtggcca	gtcctcacia	aacacagcct	attgtggaga	tcctgttaaa	aaatcagccc	1020
aaactcattg	agtttctgag	cagcttccaa	aaagaaagga	cggatgatga	gcagttcgct	1080
gacgagaaga	actacttgat	taaacagatc	cgagacttga	agaaaacggc	cccttgaaga	1140

```

gctccccggc cctgtcaca gtcagtcgtc tcatttgtcc agtttgtaca gtgtgtcatt 1200
tcagaaagtc atcattcttg ggaagacttt ggaggtgcct attttttctg ctgtaattgt 1260
tctgggtaga tggagtataa acatttgaat ggaaaaaaat taacctagaa taatatattc 1320
attttagtca aaaaaaaaaa aaaa                                     1344

```

<210> 3

<211> 341

<212> PRT

<213> Mus sp.

<300>

<308> GenBank ID No: g262934

<400> 3

```

Met Pro Phe Pro Phe Gly Lys Ser His Lys Ser Pro Ala Asp Ile
1          5          10          15
Val Lys Asn Leu Lys Glu Ser Met Ala Val Leu Glu Lys Gln Asp
20          25          30
Ile Ser Asp Lys Lys Ala Glu Lys Ala Thr Glu Glu Val Ser Lys
35          40          45
Asn Leu Val Ala Met Lys Glu Ile Leu Tyr Gly Thr Asn Glu Lys
50          55          60
Glu Pro Gln Thr Glu Ala Val Ala Gln Leu Ala Gln Glu Leu Tyr
65          70          75
Asn Ser Gly Leu Leu Gly Thr Leu Val Ala Asp Leu Gln Leu Ile
80          85          90
Asp Phe Glu Gly Lys Lys Asp Val Ala Gln Ile Phe Asn Asn Ile
95          100         105
Leu Arg Arg Gln Ile Gly Thr Arg Thr Pro Thr Val Glu Tyr Ile
110         115         120
Cys Thr Gln Gln Asn Ile Leu Phe Met Leu Leu Lys Gly Tyr Glu
125         130         135
Ser Pro Glu Ile Ala Leu Asn Cys Gly Ile Met Leu Arg Glu Cys
140         145         150
Ile Arg His Glu Pro Leu Ala Lys Ile Ile Leu Trp Ser Glu Gln
155         160         165
Phe Tyr Asp Phe Phe Arg Tyr Val Glu Met Ser Thr Phe Asp Ile
170         175         180
Ala Ser Asp Ala Phe Ala Thr Phe Lys Asp Leu Leu Thr Arg His
185         190         195
Lys Leu Leu Ser Ala Glu Phe Leu Glu Gln His Tyr Asp Arg Phe
200         205         210
Phe Ser Glu Tyr Glu Lys Leu Leu His Ser Glu Asn Tyr Val Thr
215         220         225
Lys Arg Gln Ser Leu Lys Leu Leu Gly Glu Leu Leu Leu Asp Arg
230         235         240
His Asn Phe Thr Ile Met Thr Lys Tyr Ile Ser Lys Pro Glu Asn
245         250         255
Leu Lys Leu Met Met Asn Leu Leu Arg Asp Lys Ser Arg Asn Ile
260         265         270
Gln Phe Glu Ala Phe His Val Phe Lys Val Phe Val Ala Asn Pro
275         280         285
Asn Lys Thr Gln Pro Ile Leu Asp Ile Leu Leu Lys Asn Gln Thr

```

290	295	300
Lys Leu Ile Glu Phe Leu Ser Lys Phe	Gln Asn Asp Arg Thr Glu	
305	310	315
Asp Glu Gln Phe Asn Asp Glu Lys Thr	Tyr Leu Val Lys Gln Ile	
320	325	330
Arg Asn Leu Lys Arg Ala Ala Gln Gln	Glu Ala	
335	340	

<210> 4

<211> 339

<212> PRT

<213> *Drosophila melanogaster*

<300>

<308> GenBank ID No: g1794137

<400> 4

Met Pro Leu Phe Gly Lys Ser Gln Lys Ser Pro Val Glu Leu Val		
1	5	10
Lys Ser Leu Lys Glu Ala Ile Asn Ala Leu Glu Ala Gly Asp Arg		15
20	25	30
Lys Val Glu Lys Ala Gln Glu Asp Val Ser Lys Asn Leu Val Ser		
35	40	45
Ile Lys Asn Met Leu His Gly Ser Ser Asp Ala Glu Pro Pro Ala		
50	55	60
Asp Tyr Val Val Ala Gln Leu Ser Gln Glu Leu Tyr Asn Ser Asn		
65	70	75
Leu Leu Leu Leu Leu Ile Gln Asn Leu His Arg Ile Asp Phe Glu		
80	85	90
Gly Lys Lys His Val Ala Leu Ile Phe Asn Asn Leu Leu Arg Arg		
95	100	105
Gln Ile Gly Thr Arg Ser Pro Thr Val Glu Tyr Ile Cys Thr Lys		
110	115	120
Pro Glu Ile Leu Phe Thr Leu Met Ala Gly Tyr Glu Asp Ala His		
125	130	135
Pro Glu Ile Ala Leu Asn Ser Gly Thr Met Leu Arg Glu Cys Ala		
140	145	150
Arg Tyr Glu Ala Leu Ala Lys Ile Met Leu His Ser Asp Glu Phe		
155	160	165
Phe Lys Phe Phe Arg Tyr Val Glu Val Ser Thr Phe Asp Ile Ala		
170	175	180
Ser Asp Ala Phe Ser Thr Phe Lys Glu Leu Leu Thr Arg His Lys		
185	190	195
Leu Leu Cys Ala Glu Phe Leu Asp Ala Asn Tyr Asp Lys Phe Phe		
200	205	210
Ser Gln His Tyr Gln Arg Leu Leu Asn Ser Glu Asn Tyr Val Thr		
215	220	225
Arg Arg Gln Ser Leu Lys Leu Leu Gly Glu Leu Leu Leu Asp Arg		
230	235	240
His Asn Phe Thr Val Met Thr Arg Tyr Ile Ser Glu Pro Glu Asn		
245	250	255
Leu Lys Leu Met Met Asn Met Leu Lys Glu Lys Ser Arg Asn Ile		
260	265	270

Gln Phe Glu Ala Phe His Val Phe Lys Val Phe Val Ala Asn Pro
 275 280 285
 Asn Lys Pro Lys Pro Ile Leu Asp Ile Leu Leu Arg Asn Gln Thr
 290 295 300
 Lys Leu Val Asp Phe Leu Thr Asn Phe His Thr Asp Arg Ser Glu
 305 310 315
 Asp Glu Gln Phe Asn Asp Glu Lys Ala Tyr Leu Ile Lys Gln Ile
 320 325 330
 Lys Glu Leu Lys Pro Leu Pro Glu Ala
 335

<210> 5

<211> 377

<212> PRT

<213> Caenorhabditis elegans

<300>

<308> GenBank ID No: g1255838

<400> 5

Met Pro Leu Leu Phe Gly Lys Ser His Lys Ser Pro Ala Asp Val
 1 5 10 15
 Val Lys Thr Leu Arg Glu Val Leu Thr Ile Leu Asp Lys Leu Pro
 20 25 30
 Pro Pro Lys Leu Asp Lys Asp Gly Asn Ile Gln Ser Asp Lys Lys
 35 40 45
 Tyr Asp Lys Ala Leu Asp Glu Val Ser Lys Asn Val Ala Met Ile
 50 55 60
 Lys Ser Phe Ile Tyr Gly Asn Asp Ser Ala Glu Pro Ser Ser Glu
 65 70 75
 His Val Val Gln Val Ala Gln Leu Ala Gln Glu Val Tyr Asn Ala
 80 85 90
 Asn Ile Leu Pro Met Leu Ile Lys Met Leu Pro Lys Phe Glu Phe
 95 100 105
 Glu Cys Lys Lys Asp Val Gly Gln Ile Phe Asn Asn Leu Leu Arg
 110 115 120
 Arg Gln Ile Gly Thr Arg Ser Pro Thr Val Glu Tyr Leu Gly Ala
 125 130 135
 Arg Pro Glu Ile Leu Ile Gln Leu Val Gln Gly Tyr Ser Val Pro
 140 145 150
 Asp Ile Ala Leu Thr Cys Gly Leu Met Leu Arg Glu Ser Ile Arg
 155 160 165
 His Asp His Leu Ala Lys Ile Ile Leu Tyr Ser Asp Val Phe Tyr
 170 175 180
 Thr Phe Phe Leu Tyr Val Gln Ser Glu Val Phe Asp Ile Ser Ser
 185 190 195
 Asp Ala Phe Ser Thr Phe Lys Glu Leu Thr Thr Arg His Lys Ala
 200 205 210
 Ile Ile Ala Glu Phe Leu Asp Ser Asn Tyr Asp Thr Phe Phe Ala
 215 220 225
 Gln Tyr Gln Asn Leu Leu Asn Ser Lys Asn Tyr Val Thr Arg Arg
 230 235 240
 Gln Ser Leu Lys Leu Leu Gly Glu Leu Leu Leu Asp Arg His Asn

	245	250	255
Phe Asn Thr Met	Thr Lys Tyr Ile Ser	Asn Pro Asp Asn Leu Arg	
	260	265	270
Leu Met Met Glu	Leu Leu Arg Asp Lys	Ser Arg Asn Ile Gln Tyr	
	275	280	285
Glu Ala Phe His	Val Phe Lys Val Phe	Val Ala Asn Pro Asn Lys	
	290	295	300
Pro Lys Pro Ile	Ser Asp Ile Leu Asn	Arg Asn Arg Glu Lys Leu	
	305	310	315
Val Glu Phe Leu	Ser Glu Phe His Asn	Asp Arg Thr Asp Asp Glu	
	320	325	330
Gln Phe Asn Asp	Glu Lys Ala Tyr Leu	Ile Lys Gln Ile Gln Glu	
	335	340	345
Met Lys Ser Ser	Pro Lys Glu Ala Lys	Lys Pro Lys Ser Lys Glu	
	350	355	360
Asp Glu Asn Gln	Glu Pro Ala Gly Pro	Ser Glu Gly Pro Ser Thr	
	365	370	375
Ser Gln			

Internal Application No
PCT/US 99/27027

A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C12N15/12 C07K14/47 C12Q1/68 A61K38/17 C07K16/18				
According to International Patent Classification (IPC) or to both national classification and IPC				
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 7 C12N C07K C12Q A61K				
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched				
Electronic data base consulted during the international search (name of data base and, where practical, search terms used)				
C. DOCUMENTS CONSIDERED TO BE RELEVANT				
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.		
X	MIYAMOTO HIROSHI ET AL: "Molecular cloning of a novel mRNA sequence expressed in cleavage stage mouse embryos." MOLECULAR REPRODUCTION AND DEVELOPMENT 1993, vol. 34, no. 1, 1993, pages 1-7, XP000892775 ISSN: 1040-452X	5,10		
Y	figure 4	1-4,8		
X	NATIONAL CANCER INSTITUTE, CANCER GENOME ANATOMY PROJECT (CGAP): "Homo sapiens cDNA clone similar to M025 protein" EMBEST DATABASE ENTRY HS1182722, ACCESSION NUMBER AA278473,3 April 1997 (1997-04-03), XP002133736	5,9,10		
Y	sequence	1-4,8		
-/-				
<input checked="" type="checkbox"/> Further documents are listed in the continuation of box C. <input type="checkbox"/> Patent family members are listed in annex.				
* Special categories of cited documents : <table style="width: 100%;"> <tr> <td style="width: 50%;"> "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed </td> <td style="width: 50%;"> "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family </td> </tr> </table>			"A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family
"A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family			
Date of the actual completion of the international search <div style="text-align: center; font-weight: bold;">22 March 2000</div>		Date of mailing of the international search report <div style="text-align: center; font-weight: bold;">04/04/2000</div>		
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3018		Authorized officer <div style="text-align: center; font-weight: bold;">Espen, J</div>		

INTERNATIONAL SEARCH REPORT

Internat. Application No.

PCT/US 99/27027

C. (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	NATIONAL CANCER INSTITUTE, CANCER GENOME ANATOMY PROJECT (CGAP): "Homo sapiens cDNA clone similar to M025 protein"	5,6,9,10
Y	EMEST DATABASE ENTRY HS1183536, ACCESSION NUMBER AA279145, 3 April 1997 (1997-04-03), XP002133737 sequence	1-4,8

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 99/ 27027

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
**Remark: Although claims 19,20
are directed to a method of treatment of the human/animal
body, the search has been carried out and based on the alleged
effects of the compound/composition.**
2. ☒ Claims Nos.: 17,18
because they relate to parts of the International Application that do not comply with the prescribed requirements to such
an extent that no meaningful International Search can be carried out, specifically:
See FURTHER INFORMATION sheet PCT/ISA/210
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this International application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all
searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment
of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report
covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is
restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 17,18

Claim 17 refers to an agonist of the polypeptide of claim 1 and claim 18 refers to an antagonist of the polypeptide of claim 1 without giving a true technical characterization. Moreover, no such specific compounds are defined in the application. In consequence, the scope of said claim is ambiguous and vague, and its subject-matter is not sufficiently disclosed and supported (Art. 5 and 6 PCT).

No search can be carried out for such purely speculative claims whose wording is, in fact, a mere recitation of the result to be achieved.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.